GESLIC: Genetic and Evolutionary-Based Short-Length Iris Codes

Brandon Baker, Kelvin Bryant, Gerry Dozier,

Computer Science, North Carolina Agricultural and Technical State University
1601 E. Market St., Greensboro, NC, 27411
bdbaker@ncat.edu
ksbryant@ncat.edu
gydozier@ncat.edu

Abstract—The discriminative potential of the human iris has been the center of much attention in the past decade. Feature selection has proven to be an effective means of increasing performance of current iris recognition systems. This paper introduces a novel method of reducing the size of iris codes by the use of genetic & evolutionary computing to eliminate bits corresponding to entire rings of the iris.

Keywords—Biometrics, Genetic & Evolutionary Computing, Feature Selection, Short-Length Iris Code

I. INTRODUCTION

An extraordinary potential for identification exists in the texture patterns of the iris [1]. In an effort to take advantage of this discriminative potential, a method introduced by Daugman [2] has been used to extract features from the iris and store them in an iris code. In a typical Daugman based method, a high resolution image of an iris is obtained. This image is then segmented, unwrapped and fed to a Gabor filter [2]. The image’s response to the filter is what is used as an iris code. This method simply takes the texture of the iris captured in the high resolution image and processes it in such a way that the iris texture can be represented in a binary array.

The term short-length iris codes (SLICs) was first introduced by Gentile et al. [3]. They were able to significantly reduce the number of bits needed for recognition by eliminating bits in an iris code that corresponded to entire rings of the iris. This was achieved by using a statistical technique called Kolmogorov-Smirnov to locate a region of interest within the iris. They then proceeded to build a correlation matrix which they used to determine an adequate sampling rate within a region of interest. This method yielded a 12.8x reduction in iris code size on the Multimedia University (MMU) dataset [3]. Gentile et al. then proposes a two stage iris recognition system that uses SLICs as an index into a gallery of full-length iris codes (FLICs) [4]. Such a method takes advantage of the significant size reduction of SLICs to directly reduce the computational complexity needed to perform iris recognition.

This reduction in iris code size is commonly referred to as feature selection. The order of magnitude in which the iris code is reduced in size can in turn reduce the computational load (i.e. the number of bit operations) necessary for performing iris recognition. This is the basis for the recent push to perform feature selection for iris recognition systems. In this paper, we explore the potential of using genetic and evolutionary computing (GEC) to eliminate bits corresponding to entire rings of the iris.

Hollingsworth et al. approached feature selection by removing bits that were found to be fragile (inconsistent) [5]. She simply removed bits that were consistent 60% of the time in a set of images of the same subject. Dozier et al. extends Hollingsworth’s bit fragility methodology in [6]. He successfully removed 90% of all bits that were 100% consistent and had 100% coverage across a set of images for the same subject without drastically decreasing performance.

The remainder of this paper is as follows. In Section 2, iris recognition, genetic algorithms, and how they are applied is discussed. In Section 3, the data set used to assess our method as well as the GEC parameters are discussed. In Section 4, we explore the reduction the genetic and evolutionary computing algorithms were able to obtain. In Section 5, a brief discussion of GEC resultant post processing is addressed. In Section 6, final thoughts and conclusion are discussed.

II. GENETIC & EVOLUTIONARY SHORT-LENGTH IRIS CODES

Iris recognition is typically achieved by comparing a probe (newly acquired image) to every instance in a gallery (a set of already enrolled images) [2]. A probe image is said to match an instance in the gallery with the smallest distance (hamming ratio) that is below some arbitrary threshold, typically 0.333.

We define a candidate solution (CS) as a binary vector of size 61 (equal to the number of iris rings represented in the iris code). Each CS has a value associated with it that is commonly referred to as the fitness, which is a measure of how “good” it is. Formula 1 illustrates the fitness function used to measure our CSs. This evaluation function allows our GECs to simultaneously minimize false accepts (E_FAR), false rejects (E_FRR), and the number of rings (size) in the iris code (C).

\[
fitness = \left( \frac{E_{FAR} + E_{FRR}}{C} \right) \times 10 + \left( \frac{C}{61} \right)
\]

Formula 1. Candidate solution evaluation function

We employ the use of two types of GEC, a Steady State Genetic Algorithm (SSGA) [7] and an Estimation of Distribution Algorithm (EDA) [8]. A description of how each of their initialization, selection, procreation, and replacement procedures is implemented is described below.
For an SSGA the initialization procedure is typically comprised of randomly generating an initial population of $P$ candidate solutions; these individuals are then evaluated. Parents can be selected by several means; we used binary tournament selection [7]. Procreation entailed the use of both crossover as well as a mutation. Uniform crossover [7] was used; each child was then mutated with a Gaussian mutation applied to each gene in its chromosome. The Gaussian distribution had mean of 0 and standard deviation of 1. This new child is then evaluated and replaces the current worst individual in the population.

The initialization procedure for the EDA also entailed the generation of $P$ random candidate solutions. Procreation is achieved as follows. The top 50% of the individuals in the population are selected. With this set of candidate solutions the mean ($\mu$) and standard deviation ($\sigma$) of each gene (j) in their chromosomes is calculated. These calculations are used to create an entirely new population. A new candidate solution ($x$) is created (Formula 2) by randomly generating genes using a Gaussian distribution ($N$) with a mean of 0 and standard deviation of 1. This value is then multiplied by the standard deviation and summed with the mean calculated from the elites of the current population as mentioned previously.

\[ x_j = \mu_j + \sigma_j \times N(0,1) \]

Formula 2. Generation of new candidate solution for EDA

### III. Experiments

The ICE 2006 dataset [9] was used to test our GECs. These iris codes were produced in a manner described by Thornton et al. [10]. The size of the iris codes in this data set is described as follows: 360 (degrees) x 61 (rings) x 2 (real and imaginary bits), totalling 43,920 bits. These iris codes are the result of the one-dimensional log Gabor filter that was used to produce them. This data set consists of 1,496 iris codes from 120 subjects; of which a set of 60 subjects was used to train, while the remaining 60 served as our test set. The noted baseline performance (i.e. the use of all 61 rings) yielded an accuracy of 0.9989.

### IV. Results

The GEC algorithms were observed for 10 runs each, a maximum of 1000 function evaluations, and used a population size of 20. The SSGA was implemented with binary tournament selection, a mutation rate of 1.0 and a mutation amount of 0.2. Lastly the EDA was programmed with 5 elites to run the evolutionary process.

Table 1 introduces a performance summary of the GECs. The average number of rings and accuracy are given in the second and third columns respectively. The regions of interest explored by each GEC are depicted in the fourth column. A region of interest is defined as the area between the first and last rings to appear in more than 0.5 of the runs for each GEC (i.e. 5 out of the 10 runs).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Avg. # of Rings</th>
<th>Average Accuracy</th>
<th>Interested Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDA</td>
<td>18.8</td>
<td>0.9994</td>
<td>[27, 43]</td>
</tr>
<tr>
<td>SSGA</td>
<td>15.3</td>
<td>0.9994</td>
<td>[24, 45]</td>
</tr>
</tbody>
</table>

Figure 1 shows the area explored by the GEC algorithms. The y axis is the number of times (out of the 10 runs) that a particular ring (the x axis) was used in a GESLIC. Ring 1 corresponds to the ring closest to the pupil while 61 is nearest the sclera. Rings 24 to 45 had a tendency to be explored more by both the SSGA and EDA suggesting a higher discriminatory potential than the remaining regions of the iris in this dataset. The SSGA, at best, was able to reduce from 61 to 13 iris rings; a 4.69x reduction, with no decrease in accuracy over the baseline. This reduction was obtained while maintaining an increased accuracy of 0.9993 on the test set.

The GECs were compared via a t-test on their accuracy as well as their reduction of rings. There was no statistical significance between the performances of the two algorithms in terms of accuracy; however there was a statistical difference in terms of their abilities to reduce rings. Based on statistical analysis the SSGA is the clear winner.

![Fig. 1 Record of occurrence of each ring explored by respective GECs](image-url)
V. DISCUSSION

The GESLIC reductions varied very little with respect to accuracy. Our evaluation function assumed accuracy would be a driving force and more prominent to change as size was reduced. This was not the case and because of that optimal size reduction was unobtainable by the GECs themselves. There appears to be significantly more reduction that can take place without severely reducing accuracy.

In Figure 2, we show the results of continued manual reduction within the GECs individual regions of interest. Sample rates of 1 to 10 were taken between the first and last rings within the respective regions of interest. The reduction suggested by the EDA at first glance appears to be inferior to that of the SSGA when the average rings used is considered. Once the region of interest explored by the EDA was examined we found not only was it smaller in range, but it also proved to be more resilient to even the scarcest sampling.

The sampling curve for the EDA’s region of interest shows no significant drops with lesser rings being used. There even seems to be an increase in performance from sampling rates of 9 to 10 for the EDA. A sampling rate of 10 with its region (rings 27 to 43) yielded only 2 rings used for identification and an accuracy of 0.9993 was achieved.

VI. CONCLUSION

Iris code size reduction by eliminating bits corresponding to entire rings of the iris has been successfully accomplished by way of statistical means in [3]. We employed the use of genetic and evolutionary computing in an effort to evolve short-length iris codes. Even with our implementation of an evaluation function that did not place a strong emphasis on ring reduction, a substantial number of rings were nevertheless eliminated. The GECs were able to define individual regions of interest in which we continued to reduce by modifying sampling rates. The SSGA alone was able to obtain a 4.69x reduction in iris code size and an accuracy of 0.9998 and bested that of the baseline test. Sampling within the region discovered by the EDA resulted in a 30.5x reduction. This reduction still maintained an accuracy of 0.9993 on the test set.

ACKNOWLEDGMENT

This research was funded by the Office of the Director of National Intelligence (ODNI), Center for Academic Excellence (CAE) for the multi-university Center for Advanced Studies in Identity Sciences (CASIS) and by the National Science Foundation (NSF) Science & Technology Center: Bio/computational Evolution in Action CONsortium (BEACON). The authors would like to thank the ODNI and the NSF for their support of this research.

REFERENCES


